



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/034,542	12/29/2001	Senaka Balasuriya	33692.01.0051	1424

23418 7590 08/20/2002

VEDDER PRICE KAUFMAN & KAMMHOLZ
222 N LASALLE STREET
CHICAGO, IL 60601

EXAMINER

HARPER, VINCENT P

ART UNIT	PAPER NUMBER
----------	--------------

2654

DATE MAILED: 08/20/2002

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

10/034,542

Applicant(s)

BALASURIYA, SENAKA

Examiner

V. Paul Harper

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☐ Responsive to communication(s) filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-22 is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) ____ is/are allowed.
- 6) ☒ Claim(s) 1-22 is/are rejected.
- 7) ☐ Claim(s) ____ is/are objected to.
- 8) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on ____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) ☐ The proposed drawing correction filed on ____ is: a) ☐ approved b) ☐ disapproved by the Examiner.
- If approved, corrected drawings are required in reply to this Office action.
- 12) ☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

- 13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. ____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- * See the attached detailed Office action for a list of the certified copies not received.
- 14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).
- a) ☐ The translation of the foreign language provisional application has been received.
- 15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) 2.
- 4) ☐ Interview Summary (PTO-413) Paper No(s). ____.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other:

DETAILED ACTION

Information Disclosure Statement

1. The Examiner has considered the references listed in the Information Disclosure Statement dated 12/29/01. A copy of the Information Disclosure Statement is attached to this office action.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

2. Claims 1, 2, 4-6, and 12-18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Baruch et al. (U.S. Patent Application Publication 2002/0091518 A1), hereinafter referred to as Baruch, in view of Lai et al. (U.S. Patent 6,006,183), hereinafter referred to as Lai.

Regarding claim 1, Baruch discloses a voice control system with multiple speech recognition engines. Baruch's system includes the ability to input a voice command to two recognition engines (abstract, ¶3, ¶10), which corresponds to "providing an audio command to a first speech recognition engine and at least one second speech recognition engine"; and to recognize the command with both recognition engines

generating recognition results (¶9), which corresponds to “recognizing the audio command within the first speech recognition engine to generate at least one first recognized audio command, . . . ; and recognizing the audio command within the at least one second speech recognition engine, independent of recognizing the audio command by the first speech recognition engine, to generate at least one second recognized audio command,” But Baruch does not specifically indicate that the two recognizers generate confidence values associated with their individual recognition results. However, the examiner contends that this concept was well known in the art, as taught by Lai.

In the same field of endeavor, Lai indicates the production of a score (confidence level) for each word that is recognized (col. 2, lines 61-63).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch by specifically providing a confidence level with each recognition result, as taught by Lai, for the purpose of determining the degree of confidence associated with a given recognition event.

Regarding claim 2, Baruch in view of Lai teach everything claimed, as applied above (see claim 1); in addition, Baruch teaches the choosing between the first recognized result of the first recognition engine and a second recognized result of the second engine (¶9), which corresponds to “selecting at least one recognized audio command having a recognized audio command confidence value from the at least one first recognized audio command and the at least one second recognized audio

command based on the at least one first confidence value and the at least one second confidence value.”

Regarding claim 4, Baruch in view of Lai teach everything claimed, as applied above (see claim 2); in addition, Baruch teaches the choosing of the command that in a control system such as Baruch’s would inherently be executed (§9), which corresponds to “executing at least one operation based on the at least one recognized audio command.”

Regarding claim 5, Baruch in view of Lai teach everything claimed, as applied above (see claim 2); in addition, Baruch teaches that the voice controlled apparatus can give user feedback (§7), which corresponds to “verifying the at least one recognized audio command.”

Regarding claim 6, Baruch in view of Lai teach everything claimed, as applied above (see claim 1); in addition, Baruch teaches that if a voice input is not recognized the system may provide a visual and/or audible message (§40), which corresponds to “generating an error notification.” But Baruch in view of Lai do not specifically teach that this would occur “when the at least one first confidence value and the at least one second confidence value are below a minimum confidence level.” However it is inherent in a system such as Baruch in view of Lai’s where a recognition decision is made based on confidence levels that if the results of both recognition units are below the respective minimum confidence levels, an error would result.

Regarding claim 12, Baruch discloses a voice control system with multiple speech recognition engines. Baruch’s system includes the ability to input a command

Art Unit: 2654

from a microphone to a recognition engine (abstract, ¶3, ¶10), which corresponds to “a first speech recognition means, operably coupled to an audio subsystem, for receiving an audio command and generating at least one first recognized audio command.” But Baruch does not specifically indicate “the at least one first recognized audio command has a first confidence value.” However, the examiner contends that this concept was well known in the art, as taught by Lai.

In the same field of endeavor, Lai indicates the production of a score (confidence level) for each word that is recognized (col. 2, lines 61-63).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch by specifically providing a confidence level with each recognition result, as taught by Lai, for the purpose of determining the degree of confidence associated with a given recognition event.

Baruch's system includes the ability to input a command from a microphone to a second recognition engine (abstract, ¶3, ¶10, Fig. 1), which corresponds to “a second speech recognition means, operably coupled to the audio subsystem, for receiving the audio command and generating, independent of the first speech recognition means, at least one second recognized audio command,” But Baruch does not specifically indicate “each of the at least one second recognized audio command has a second confidence value.” However, the examiner contends that this concept was well known in the art, as taught by Lai.

In the same field of endeavor, Lai indicates the production of a score (confidence level) for each word that is recognized (col. 2, lines 61-63).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch by specifically providing a confidence level with each recognition result, as taught by Lai, for the purpose of determining the degree of confidence associated with a given recognition event.

Baruch's system has a control unit **18** connected directly to the recognition engines (**¶23**) and the ability to choose between the recognition results from the first and second recognizers (**¶9**, Fig. 1), which corresponds to "a means, operably coupled to the first speech recognition means and the second speech recognition means, for receiving the at least one first recognized audio command and the at least one second recognized audio command."

Regarding claim 13, Baruch in view of Lai teach everything claimed, as applied above (see claim 12). In addition, Baruch teaches: the choosing between a first recognized result of the first engine and a second recognized result of the second engine where the recognition units are coupled to a control unit (**¶9**, Fig. 1, **18**), which corresponds to "a dialog manager operably coupled to the means for receiving, wherein the means for receiving selects at least one recognized audio command having a recognized confidence value from the at least one first recognized audio command and the at least one second recognized audio command based on the at least one first confidence value and the at least one second confidence value"; and that confirmation may be required at any point before proceeding and may require verbal or visual prompting (**¶50**), which corresponds to "the selected at least one recognized audio command is provided to the dialog manager."

Regarding claim 14, Baruch in view of Lai teach everything claimed, as applied above (see claim 12); in addition, Baruch teaches: that the recognition results go to the control unit 18 and a choice between a first recognized result of the first engine and a second recognized result of the second engine is made (§9), which corresponds to “dialog manager determines a dialog manager audio command from the at least one recognized audio command based on the at least one recognized audio command confidence levels”; and the result of the speech recognition is used in command and control applications (§3), which corresponds to “and wherein the dialog manager executes an operation in response to the dialog manager audio command.”

Regarding claim 15, Baruch in view of Lai teach everything claimed, as applied above (see claim 12). In addition, Baruch teaches that through voice commands a user can access a list of previously selected languages where the list may be provided over a loudspeaker (§44), which corresponds to “wherein the dialog manager accesses a content server and retrieves encoded information in response to the dialog manager audio command.”

Regarding claim 16, Baruch in view of Lai teach everything claimed, as applied above (see claim 15). In addition, Baruch teaches that a list of requested languages may be provided by loudspeaker (§44), which corresponds to “a speech synthesis engine operably coupled to the dialog manager, wherein the speech synthesis engine receives speech encoded information from the dialog manager and generates speech formatted information.”

Regarding claim 17, Baruch in view of Lai teach everything claimed, as applied above (see claim 16). In addition, Baruch teaches that a speaker **34** is attached to a digital communication unit **30** and a control unit **18**, and that this subsystem can generate audio prompts ($\P 41$), which corresponds to "the audio subsystem is operably coupled to the speech synthesis engine, wherein the audio subsystem receives the speech formatted information and provides an output message."

Regarding claim 18, Baruch in view of Lai teach everything claimed, as applied above (see claim 17). In addition, Baruch teaches that if the input is not recognized an audible message may be given ($\P 41$), which corresponds to "when the comparator provides the dialog manager with an error notification, the output message is an error statement."

3. Claims 3, 7-11, and 19-22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Baruch in view of Lai, and further in view of Baker (U.S. Patent No. 6,122,613).

Regarding claim 3, Baruch in view of Lai teach everything claimed, as applied above (see claim 2), but Baruch in view of Lai do not specifically teach "weighting the at least one first confidence value by a first weight factor and weighting the at least one second confidence values by a second weight factor." However, the examiner contends that this concept was well known in the art, as taught by Baker.

In the same field of endeavor, Baker teaches the combining of the recognition results based on the weighting factors (col. 3, 38-42).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch in view of Lai by specifically weighing the results from each recognizer, as taught by Baker, for the purpose of assigning a greater weight to the recognizer known to be more accurate (Baker, col. 3, line 42).

Regarding claim 7, Baruch discloses a voice control system with multiple speech recognition engines **10**. Baruch's system includes the ability to input an audio command to two recognition engines (abstract, ¶3, ¶10), which corresponds to "providing an audio command to a terminal speech recognition engine and at least one . . . [additional] speech recognition engine; recognizing the audio command within the terminal speech recognition engine to generate at least one terminal recognized audio command." But Baruch does not specifically teach "wherein the at least one terminal [or network] recognized audio command has a corresponding terminal confidence value." However, the examiner contends that this concept was well known in the art, as taught by Lai.

In the same field of endeavor, Lai indicates the production of a score (confidence level) for each word that is recognized (col. 2, lines 61-63).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch by specifically providing a confidence level with each recognition result, as taught by Lai, for the purpose of determining the degree of confidence associated with a given recognition event.

In addition, Baruch does not specifically teach that the second recognizer is a network speech recognition engine. However, the examiner contends that the concept

of the use of a second recognizer connected on a network was well known in the art, as taught by Baker.

In the same field of endeavor, Barker teaches speech recognition using two recognizers applied to the same input sample, where the second recognizer can be a network device (Fig. 3, **315, 309**).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch by specifically providing an additional recognizer accessed by a network connection, as taught by Baker, for the purpose of providing access to a more powerful recognizer through a network connection.

In addition, Baruch does not specifically teach the step of "recognizing the audio command within the at least one network speech recognition engine to generate at least one network recognized audio command, wherein the at least one network recognized audio command has a corresponding network confidence value." However, the examiner contends that this concept was well known in the art, as taught by Baker.

Baker further teaches that the output of the network recognizer is assigned a score (or confidence level) (abstract).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch in view of Baker, as taught by Baker, to aid in the decision process when selecting from between the recognition candidates.

In addition, Baruch in view of Lai and Baker teach the choosing between a first recognized result of the first engine and a second recognized result of the second engine (Baruch, ¶19) where when confidence values are used, as taught above by Lai

and Baker, these values would be used in the recognition selection process, which corresponds to "selecting at least one recognized audio command having a recognized audio command confidence value from the at least one terminal recognized audio command and the at least one network recognized audio command."

Regarding claim 8, Baruch in view of Lai and Baker teach everything claimed, as applied above (see claim 7); in addition, Baruch teaches that if a voice input is not recognized the system may provide a visual and/or audible message (¶40), which corresponds to "generating an error notification." But Baruch in view of Lai and Baker do not specifically teach that this would occur "when the at least one terminal confidence value and the at least one network confidence value are below a minimum confidence level." However it is inherent in a system such as Baruch in view of Lai's where a recognition decision is made based on confidence levels that if the results of both recognition units are below the respective minimum confidence levels, an error would result.

Regarding claim 9, Baruch in view of Lai and Baker teach everything claimed, as applied above (see claim 7), but Baruch in view of Lai and Baker do not specifically teach "weighting the at least one terminal confidence value by a terminal weight factor and the at least one network confidence value by a network weight factor." However, the examiner contends that this concept was well known in the art, as taught by Baker.

Baker further teaches the combining of the recognition results based on the weighting factors (col. 3, 38-42).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch in view of Lai and Baker by specifically weighing the results from each recognizer, as taught by Baker, for the purpose of assigning a greater weight to the recognizer known to be more accurate (Baker, col. 3, line 42).

Regarding claim 10, Baruch in view of Lai and Baker teach everything claimed, as applied above (see claim 7) including the assignment of a confidence level to the recognition events (Lai, col. 2, lines 61-63; Baker, abstract), but Baruch in view of Lai and Baker do not specifically teach "filtering the at least one recognized audio command based on the at least one recognized audio command confidence value." However, the examiner contends that this concept was well known in the art, as taught by Lai.

Lai further discloses the ability to select score thresholds above or below which recognized words are displayed (col. 3, lines 36-40)

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch in view of Lai and Baker by specifically supporting the filter capability, as taught by Lai, for the purpose of determining what the minimum confidence level for recognition will be.

In addition, Baruch teaches the choosing of a command based on the results from the recognizers (¶19) where the choice would inherently be the command with the highest confidence, and in a control system such as Baruch's the recognized command would inherently be executed, which corresponds to "executing an operation based on

the recognized audio command having the highest recognized audio command confidence value.”

Regarding claim 11, Baruch in view of Lai and Baker teach everything claimed, as applied above (see claim 7); in addition Baruch teaches the ability of the system to get confirmation from the user (¶50), which corresponds to “verifying the at least one recognized audio command to generate a verified recognized audio command”; and in a control system such as Baruch’s the execution of the command would inherently follow the affirmation, which corresponds to “executing an operation based on the verified recognized audio command.”

Regarding claim 19, Baruch discloses a voice control system with multiple speech recognition engines **10**. Baruch’s system includes the ability to input an audio command into a microphone **12** connected to a recognition engine (abstract, ¶3, ¶10), which corresponds to “a terminal speech recognition engine operably coupled to a microphone and coupled to receive an audio command and generate at least one terminal recognized audio command.” But Baruch does not specifically teach “wherein the at least one terminal recognized audio command has a corresponding terminal confidence value.” However, the examiner contends that this concept was well known in the art, as taught by Lai.

In the same field of endeavor, Lai indicates the production of a score (confidence level) for each word that is recognized (col. 2, lines 61-63).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch by specifically providing a confidence

level with each recognition result, as taught by Lai, for the purpose of determining the degree of confidence associated with a given recognition event.

In addition, Baruch does not specifically disclose "at least one network speech recognition engine operably coupled to the microphone and coupled to receive the audio command and generate at least one network recognized audio command, independent of the terminal speech recognition engine, wherein the at least one network recognized audio command has a corresponding network confidence value." However, the examiner contends that the concept of the use of a second recognizer connected on a network was well known in the art, as taught by Baker.

In the same field of endeavor, Barker teaches speech recognition using multiple recognizers applied to the same input sample, where the second recognizer can be a network device and that a confidence value is associated recognition candidates (abstract, Fig. 3, **315, 309**).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Baruch by specifically providing an additional recognizer accessed by a network connection, as taught by Baker, for the purpose of providing access to a more powerful recognizer through a network connection.

In addition, Baruch teaches the connecting of the recognition engines to a control unit where the results are send **18**, which corresponds to "a comparator operably coupled to the terminal speech recognition engine operably coupled to receive the at least one terminal recognized audio command and further operably coupled to the at

Art Unit: 2654

least one network speech recognition engine operably coupled to receive the at least one network recognized audio command.”

In addition, Baruch in view of Lai and Baker disclose the choosing between a first recognized result of the first engine and a second recognized result of the second engine (Baruch, ¶9) and where confidence values are used in the recognition selection process, as taught above by Lai and Baker, which corresponds to “a dialog manager operably coupled to the comparator, wherein the comparator selects at least one recognized audio command having a recognized confidence value from the at least one terminal recognized audio command and the at least one network recognized audio command based on the at least one terminal confidence value and the at least one network confidence value.”

Baruch also teaches that the system may require confirmation before proceeding (¶50), which corresponds to “the selected at least one recognized audio command is provided to the dialog manager.”

Regarding claim 20, Baruch in view of Lai and Baker teach everything claimed, as applied above (see claim 19). In addition, Baruch discloses a voice control system where a choice is made between the recognition results of two recognizers with a possible confirmation before execution of the command (abstract, ¶9, ¶50), which corresponds to “the dialog manager determines a dialog manager audio command from the at least one recognized audio commands based on the at least one recognized audio command confidence levels and wherein the dialog manager executes an operation in response to the dialog manager audio command.”

Regarding claim 21, Baruch in view of Lai and Baker teach everything claimed, as applied above (see claim 20). In addition, Baruch discloses a technique where a user may call up a list of languages and then choose one by uttering its name (§44), which corresponds to "wherein the dialog manager accesses a content server and retrieves encoded information in response to the dialog manager audio command."

Regarding claim 22, Baruch in view of Lai and Baker teach everything claimed, as applied above (see claim 21). In addition, Baruch discloses a loudspeaker for audible output messages that is connected to the control unit through the digital communications unit (Fig. 1, §41), which corresponds to "wherein the speech synthesis engine receives speech encoded information from the dialog manager and generates speech formatted information; and a speaker operably coupled to the speech synthesis engine, wherein the speaker receives the speech formatted information and provides an output message."

Citation of Pertinent Art

4. The following prior art made of record but not relied upon is considered pertinent to the applicant's disclosure:

a) Fiscus et al., "A post-processing system to yield reduced word error rates:

Recognizer Output Voting Error Reduction (ROVER)," Proceedings of 1997 IEEE Workshop on Automatic Speech Recognition and Understanding, Dec. 1997, describes a system with multiple automatic speech recognition units that uses a voting process to select the output.

Art Unit: 2654

- b) Brown et al., (U.S. Patent No. 6,377,922) describes a distributed recognition system using multiple recognizers.
- c) Hunt et al., (U.S. Patent No. 5,365,574) describes a speech recognition system with adjustable thresholds.
- d) Modi et al., (U.S. Patent No. 6,125,345) describes a speech recognition system with multiple confidences measures.
- e) Reformato et al., (U.S. Patent Application Publication No. 2001/0036255) describes a multiple speech recognition system where the recognizers have differing capabilities.

Conclusion

Any response to this office action should be mailed to:

Commissioner of Patents and Trademarks
Washington, D.C. 20231

or faxed to:

(703) 872-9314

Hand-delivered responses should be brought to:

Crystal Park II
2121 Crystal Drive
Arlington, VA.
Sixth Floor (Receptionist)

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dr. V. Paul Harper whose telephone number is (703)

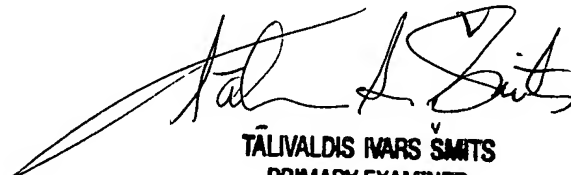
Art Unit: 2654

305-4197. The examiner can normally be reached on Monday through Friday from 8:00 a.m. to 4:30 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Marsha D. Banks-Harold, can be reached on (703) 305-4379. The fax phone number for the Technology Center 2600 is (703) 872-9314.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Technology Center 2600 Customer Service office whose telephone number is (703) 306-0377.

VPH/vph
August 16, 2002



TĀLIVALDIS NARS SMITS
PRIMARY EXAMINER